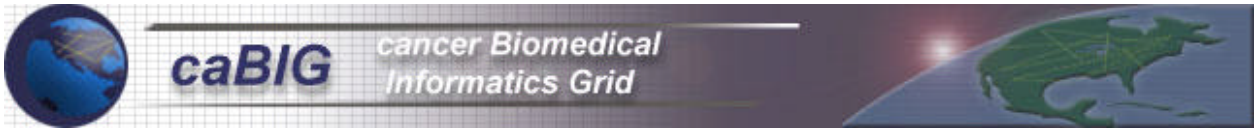

Integrative Cancer Research Workspace

ICR Workspace Working Group Teleconference Meeting Minutes

Date, Time & Location:	November 10, 2004, 2:00pm – 3:00pm EST

Attendees:

First Name	Last Name	Institution
Chris	Abajian	Fred Hutchinson Cancer Research
Margaret	Borwhat	Patient Advocate
Terry	Braun	University of Iowa
Andrea	Califano	University of Columbia
Thomas	Casavant	Universtiy of Iowa
Hong	Chan	University of Georgetown
Leo	Cheung	University of Hawaii
Robert	Clarke	University of Georgetown
Cal	Collins	Akaza Research
Tommie	Curtis	NCICB
Hong	Dang	ALPHA-GAMA Technologies
Roger	Day	UNI of Pittsburgh
Joel	Dubbels	IBM
Brian	Gilman	Cold Spring Harbor Lab
Judith	Goldberg	NYU
Ajay	Jain	UCSF
David	Jewell	Dartmouth College
David	Kane	NCI-CCR
Michael	Keller	Booz Allen Hamilton
Chris	Kinglsey	UCSF
Gene	Kraus	Karmanos Cancer Institute
Doctor	Kutbuddin	Burnham Institute
Alex	Lash	Sloan-Kettering
Ted	Liefeld	MIT-Broad
Jack	London	Jefferson
James	Lyonsweiler	University of Pittsburgh Cancer Institute
Subha	Madhavan	NCI
Arumani	Manisudaram	Booz Allen Hamilton
Steve	Marron	University of North Carolina
Jason	Miller	Incogen
Tom	Moloshok	Fox Chase Cancer Center
Rakesh	Nagarajan	Wash U
Michael	Ochs	Fox Chase Cancer Center
Harold	Riethman	Wistar Institute
John	Rux	Wistar Institute
Nitin	Sawhne	Akaza Research
Todd	Scheetz	University of Iowa
Louise	Showe	Wistar Institute
Craig	Street	University of Pennsylvania
Mark	Watson	Wash U

	Mathiew	Wiepert	Mayo Clinic
	Quing	Xiao	JPL
	Dong	Xing	UNC Capital Hill
	Liming	Yang	NCI
	Vincent	Yau	Oregon Health

Summary of outcomes from Strategic Planning Working Group face-to-face meeting

Michael Ochs gave an overview of the key outcomes at this week's face-to-face meeting of the Strategic Planning Working Group (November 8-9, 2004). Resources for this meeting will be posted in the web in the near future.

1. Principles to guide the working group:
   a. All software should be designed and engineered for change.
   b. There should be a bias toward Action. If there are multiple choices, make the best choice possible with the available information and then proceed.
   c. Three and five-year goals were discussed

Three-year goals:

- Tools have a positive impact on caBIG participants
- All projects are developed at the Gold standard of compliance
- There is interoperability and adoption of tools in caBIG

Five-year goal:

- Widespread adoption of caBIG into the larger cancer community

2. In ICR, most development is targeted at existing tools; therefore, interoperable tools should be available on the grid by the end of the year.
   a. ICR and Architecture have the most significant level of interaction in caBIG and this link is critical.
   b. Data models and use cases must be shared across ICR.
   c. There is the desire that all groups use the caDSR tools to create and reuse data elements or to directly adopt the caBIO objects. caBIO objects are stable and can be used for development.
3. Security models should be included in all development.

Q&A:

Q (Jim Lyons-Weiler): Is there going to be a workshop for the caCORE resources?

A (Michael Ochs): Yes, there was discussion of this at the meeting. There is planning for online training. Recommendation: Cancer Center Participants should be vocal about the need.

A (Arumani): Note that the Training Group is being reorganized to address such needs.

Q (Jim Lyons-Weiler): If Gold is a 3 year goal, why is having interoperable tools a 1 year goal?

A (Michael Ochs): There are grid reference implementations this year that will show interoperability on the grid. Also, one can have interoperable tools with Silver level compliance.

Summary of outcomes from Architecture, VCDE joint face-to-face meeting	Arumani Manisundaram and Mike Keller gave a summary of the outcomes of the recent Architecture and VCDE joint face-to-face meeting (October 25-26, 2004) Their presentation can be found at: **http://cabig.nci.nih.gov/workspaces/ICR/Meetings/ICR_Workspace/Teleconferences/20041110_ArchVCDESummary_presentation** All presentations from the joint face-to-face meeting can be found at: **http://cabig.nci.nih.gov/workspaces/Architecture/Meetings/f2f-meetings/ARCH-VCDE-F2F/view**
Updates from other Workspaces:	**Tissue Banks and Pathology Tools (Mark Watson)** • The TBPT WS last convened by teleconference on 11/2/04. UPMC and WU (joint developers in the WS) presented their vision for caTISSUE and general comments were solicited. • caTISSUE will integrate diverse data sources to annotate biospecimens. A "clinical annotation engine" and corresponding data mapping tools will be developed by UPMC as a system to collect and integrate clinical annotation and map it to caTISSUE objects. CaTIES will be an early deployed tool that will extract and de-identify data elements from text pathology reports. • caTISSUE will contain a core system (caTISSUE Lite) that will provide the framework object model to which additional functional modules (such as the clinical annotation engine) can be added. Institutions may adopt caTISSUE Lite if they have no existing system or their existing system is inadequate. Alternatively, institutions may elect to map their existing tissue bank system to the caTISSUE object model. CaTISSUE Lite will be developed by WU. • It was emphasize that caTISSUE should not try to reinvent the wheel and that there were many excellent systems currently functioning, from which use cases could be readily constructed. • Of particular relevance to ICR, WU will focus on experimental annotation of biospecimens (interaction with Translational SIG will be important). • The TBPT WS will hold it's first face-to-face meeting on 11/30-12/1 at UPMC. The next teleconference is scheduled for 11/16/04. **Data Sharing and Intellectual Capital (Terry Braun)** DSIC has divided into "topic teams": • Regulatory Team and • Proprietary Team Two topic teams are starting to develop outlines for position papers. From each team, a member has volunteered to be the team leader, whose contribution includes setting agendas for team meetings. - Proprietary Team Leader: Pat Harsche-Weeks - Regulatory Team Leader: Howard Bilofsky Each team will coordinate an effort to solicit use cases from other

	WS/WG. Proprietary Team: Oct 28: Develop working list of definitions: license, software, data, biospecimens Discuss the range of license/sharing standards for sharing software, data, and biospecimens. Regulatory Team: Nov 1: Goals: • Raise awareness across caBIG of importance and impact of HIPAA and other Regulations on Collaborative Cancer research • Ensure that all caBIG participants are aware of the need for all caBIG activities, services, and information resources to operate in full compliance with all applicable regulations and laws • Provide guidance to other caBIG components regarding the need to design, deploy, and operate caBIG systems in a manner that facilitates regulatory compliance • Create foundation for design and implementation of caBIG systems to enhance performance
Update of SIG activities	Juli gave an overview of ICR SIG activities this month: **Translational**: Terry Braun gave an overview of the TrAPSS architecture. Subha Madhavan from NCICB updated the group an emerging collaboration between NCI and NCRI (UK organization). One need is the creation of an object model for the integration of clinical and functional genomics data – this SIG may get involved in this effort. **Pathways**: Group began discussing pathways exchange standards to be used by the group. BioPAX and PSI-MI are both under discussion. Decision will be driven by project use cases. **Microarray**: UAT testing started Nov. 1 and is ongoing. A caArray Developers meeting and listserv will be established to support the many Developers in ICR that will be making use of caArray. Also had a presentation of the NCI-CCR's NCI-60 project **Genome Annotation:** The team from Burnham presented the Cancer Molecular Pages project. Gene CDE focus group has continued to discuss how to address the definition of this data type. A meeting yesterday with Denise Warzel greatly helped in clarifying a number of key questions about possible ways to proceed. We may now be in a position to revise the whitepaper Rakesh Nagarajan had originally drafted. **Data Analysis and Statistical Tools**: UCSF's Magellan project was presented at the last meeting. Team has begun discussing mechanisms for data input and data export and will benefit from the caArray developers meeting mentioned earlier. **Proteomics**: RProteomics team had their first face-to-face meeting last week at Duke. Penn and OHSU are Adopters on this project. The FCCC Proteomics LIMS team presented an overview of their project at the last SIG meeting. Face-to-face meeting with their adopter, Moffitt, in early December.
General	- Nearly all funded participants in ICR have received their task orders but we have

caBIG Reminders	not received signatures from many of the centers. Please send in your signed task orders as soon as possible. We must received the signed task orders before invoices can be processed. - Please submit monthly reports for your ICR Workspace Participant task orders. For Centers receiving their task orders in August or September, the center can start reporting for activities in August. For Centers receiving their task orders in October, the center can start reporting for activities in September. Prior to CVS and caMP being available, these should be emailed directly to Juli. The template for the report can be found at: http://cabig.nci.nih.gov/workspaces/ICR/Templates/caBIG%20Monthly%20WSWG%20Report%20template.doc/view - UML-based CDE creation training is still ongoing. The information being presented is particularly important for ICR Developers. Please see the caDSR training website for dates, times and materials: http://ncicb.nci.nih.gov/NCICB/core/caDSR/Training - As a reminder, every group should be aware that all hardware purchased under a federal government contract is officially the property of the federal government as is subject to return once the project is complete. - CVS and caMP account information has now been made available to Juli and she will make this information available for distribution by the end of the week.
General Q&A	Q (Steve Marron): Regarding the finding that XML schema may be insufficient for metadata representation – how does this affect our approach to CDE creation: A (David Kane): XML will still define the pieces to be exchanged. For example, you can still use MAGE-ML to describe data sets but you may need additional documents as well. This really pertains to Gold development. Q (Steve Marron): Does anyone know of a good tool for converting microarray data to MAGE-ML? A (Juli Klemm): The current version of caArray supports this for Affymetrix and GenePix files. A (Subha Madhavan): For the SMD format that Steve is interested in, this would have to be in a future version of caArray. One suggestion would be to convert SMD format to GenePix format – then you could use caArray. A (other): The MAGE-stk toolkit may be useful in this regard.

Action Items:

Name Responsible	Action Item	Date Due	Notes
Juli Klemm	Include links to Architecture and VCDE presentations in meeting minutes.	11/17/04	
Juli Klemm	Distribute CVS and caMP account information	11/12/04	